

# El componente de selección adversa de la horquilla de precios cotizada: una revisión de los modelos de estimación\*

## The adverse selection component for the bid-ask spread: A revision of its estimation models

JOSÉ E. FARINÓS

C. JOSÉ GARCÍA

ANA M.<sup>a</sup> IBÁÑEZ

*Departament de Finances Empresariales. Universitat de València*

Recibido en mayo de 2004; aceptado en junio de 2005

### Resumen:

*Una de las principales preocupaciones en el área de la microestructura del mercado ha sido la estimación de los componentes no observables de la horquilla de precios a partir de las series de datos que proporcionan los mercados financieros, despertando quizá un mayor interés el de selección adversa por la implicaciones que supone la existencia del mismo. Esto ha provocado el desarrollo de numerosos modelos empíricos que, basándose en las propiedades estadísticas de las series de precios, proporcionan dichas estimaciones. La mayor disponibilidad de datos existentes en los mercados ha permitido el desarrollo en los últimos años de modelos basados en técnicas estadísticas más complejas como son el método generalizado de momentos o la metodología VAR y cuya base de partida es la dinámica de la formación del precio, y, en concreto, cómo la información privada de las transacciones se recoge en los nuevos precios cotizados. El objetivo de este trabajo es analizar este último grupo de trabajos, es decir, aquellos modelos de estimación de los componentes de la horquilla basados en la dinámica de la formación de precios que, además de permitir la estimación del componente de selección adversa en series temporales, suponen una herramienta fundamental para analizar el proceso de incorporación de la información a los precios cotizados en los distintos mercados.*

### Palabras clave:

*microestructura de los mercados financieros, negociación informada, horquilla de precios, selección adversa, costes de transacción*

**Clasificación JEL:** G12, G34, D82

### Abstract:

*One of the main interests of market microstructure is the estimation of the bid-ask spread components from financial data, specially the adverse selection component given the implications of its own existence. As a*

---

\* Los autores agradecen los comentarios realizados por los evaluadores anónimos de la revista. Este trabajo ha contado con el soporte financiero de la Generalitat Valenciana proyecto GV04B210.

*result, several empirical models based on price time-series statistical properties have been developed in order to estimate them. Recent greater financial data availability has allowed the development of models that focus on price discovery and use more statistical complex methodologies like GMM or VAR. This paper analyses this set of models that allows the estimation of the bid-ask spread components from price dynamics, specifically, the estimation of the adverse selection component in time series. Actually, this sort of models are a powerful tool to investigate how information is incorporated into quotes.*

**Keywords:**

*Market microstructure, insider trading, spread, adverse selection component, transaction cost.*

**JEL classification:** G12, G34, D82

## 1. INTRODUCCIÓN

El estudio de la microestructura de los mercados financieros se puede entender como el análisis del proceso de formación del precio en los mercados teniendo en cuenta el funcionamiento y características de los mismos. La consideración para ello de los distintos tipos de agentes que actúan en los mercados y sus estrategias de negociación desemboca en una ruptura de la visión clásica tradicional Walrasiana donde se proponía la existencia de un único precio de equilibrio que vacía el mercado y en la aparición, por el contrario, de ciertas fricciones. Estas fricciones vienen recogidas fundamentalmente en el coste de transacción, entendido como el coste que supone demandar liquidez por parte del inversor dado que el precio para la compra y la venta es distinto y que, por tanto, se identifica con la horquilla de precios. Dicha horquilla viene dada por la diferencia entre los precios para la compra y para la venta cotizados por los oferentes de liquidez en los mercados dirigidos por precios y por la horquilla implícita del libro de órdenes en aquellos mercados que, como el nuestro, funcionan en base a las mejores órdenes de compra y venta introducidas en dicho libro.

A este respecto, y tras el trabajo de Demsetz (1968), surgen numerosos estudios teóricos en la línea planteada anteriormente que tratan formalmente de describir la formación del precio en determinados supuestos y que identifican los componentes de la horquilla de precios cotizadas. De estos trabajos cabe destacar la diferencia entre un primer grupo que considera la existencia de dos tipos de agentes, el oferente de inmediatez y el que opera por motivos de liquidez [Ho y Stoll (1981) y Amihud y Mendelson (1980), entre otros] y aquellos que consideran, además, agentes negociando en el mercado con información privada sobre el valor cierto de los activos [Bagehot (2071), Copeland y Galai (1983), Glosten y Milgrom (1985) y Kyle (1985, 1989), entre otros]. La diferencia fundamental entre ambos grupos reside en el hecho de que en los segundos se identifica un componente de selección adversa en los costes de transacción relacionado con el contenido informativo de las transacciones y que tiene un impacto permanente en el precio. Este último componente ha sido el que más interés ha despertado dada la importancia que el mismo tiene en temas tales como el estudio de la negociación con información privada o del proceso de formación de precios.

Por otra parte, y de acuerdo a las sugerencias de los modelos teóricos de microestructura mencionados, surgen toda una serie de trabajos empíricos con distintas propuestas

para la estimación de los diferentes componentes de la horquilla a partir de las series de datos observadas en el mercado, por lo que la investigación empírica referente a la estimación de los costes de transacción y los modelos desarrollados han estado condicionados por la cantidad y características de los datos, y han evolucionado a partir de la mayor disponibilidad de éstos.

Un primer grupo de estos trabajos basan la estimación de dichos componentes en el concepto de covarianza obteniendo estimadores en sección cruzada para todo el periodo temporal que implica la estimación de dichas covarianzas. En esta línea están los modelos de Roll (1984), Stoll (1989) y Kim y Ogden (1996) entre otros.

La mayor disponibilidad de datos en las series financieras proporcionadas por los mercados ha permitido el desarrollo de estimaciones alternativas a las propuestas por los modelos señalados. Así, surge un conjunto de trabajos que proporcionan estimadores de los distintos componentes de la horquilla, pero que para proceder a su estimación necesitan disponer de datos intradía ya que la mayoría requieren conocer si la orden ejecutada fue de compra o de venta.<sup>1</sup> En estos modelos se estima la importancia de los distintos componentes de la horquilla para cada activo de forma individualizada y, a diferencia de los modelos basados en las covarianzas, su estimación se realiza en series temporales y no en sección cruzada. En este trabajo recogemos los modelos más relevantes de este segundo grupo.

El trabajo se estructura como sigue. En primer lugar desarrollamos con detalle los distintos modelos estudiados. En segundo lugar se analizan las semejanzas y diferencias entre ellos y su aplicabilidad en el mercado bursátil español dada la disponibilidad de datos. Por último concluimos sobre los aspectos más relevantes tratados en el trabajo.

## 2. HIPÓTESIS Y DESARROLLO DE LOS DISTINTOS MODELOS

El primero de los modelos empíricos que basó su estimación de la descomposición de la horquilla en el signo de las transacciones fue el de *Glosten y Harris (1988)*. Las hipótesis de partida de dicho modelo son las mismas que las el modelo seminal planteado por Roll en 1984, excepto que incorpora la existencia de agentes con información heterogénea y considera el redondeo.<sup>2</sup>

El modelo permite estimar dos componentes de la horquilla cotizada. Un primer componente denominado transitorio que no está relacionado con el valor cierto del activo y que, además de remunerar al oferente de liquidez, recoge los costes de inventario que surgen por la aversión al riesgo, los costes de liquidación y otros costes de transacción que ha de pagar el especialista. Un segundo componente, el coste de selección adversa, que surge porque el oferente de liquidez no puede identificar si está negociando con agentes con información privada y se considera permanente ya que es debido al cambio en las expectativas sobre el valor cierto del activo ante una determi-

<sup>1</sup> Sin embargo, el modelo de Glosten y Harris (1988), que veremos a continuación, es estimado por los autores sin conocer este dato. Para ello utilizan maximaverosimilitud y hacen determinados supuestos sobre las funciones de distribución de las distintas variables.

<sup>2</sup> Véase Roll (1984).

nada transacción, por tanto no produce autocorrelación en las variaciones de los precios de transacción.

La idea del trabajo es estimar los dos componentes por separado basándose en las series de precios de transacción y en las diferentes propiedades de ambos componentes. Concretamente, se centran en el efecto permanente que el componente de selección adversa tiene en la fijación de la horquilla. De hecho, a los dos componentes que estima los denomina componente de selección adversa y componente transitorio de la horquilla. Además, y siguiendo la postura defendida por Easley y O'Hara (1987), quienes sugieren que al menos parte de la dependencia de la horquilla con el tamaño de la orden es debida a la información asimétrica, el modelo permite que el componente de selección adversa dependa del tamaño de la orden.

Los autores definen el proceso de generación para el valor cierto del activo,  $M_t$ , y para los precios de transacción sin considerar el redondeo,  $P_t$ , como se indica en las expresiones [1] y [2], respectivamente.

$$M_t = M_{t-1} + E_t + Q_t Z_t \quad [1]$$

$$P_t = M_t + Q_t C_t \quad [2]$$

En las expresiones [1] y [2],  $Q_t$  es un indicador del signo de la orden que tomará valor  $+1$  si la transacción fue iniciada por una orden de compra, es decir, si se ejecutó al mejor precio de venta, y  $-1$  si la transacción fue iniciada por una orden de venta, esto es, si se ejecutó al mejor precio de compra mientras que  $Z_t$  y  $C_t$  son, respectivamente, el componente de selección adversa y el componente transitorio de la horquilla.

El valor cierto del activo se ajusta por la llegada de nueva información entre los instantes  $t-1$  y  $t$  recogida en  $E_t$ , y por la información privada que se desprende de la transacción en  $t$ ,<sup>3</sup> ajuste que coincide con el componente de selección adversa, que provoca una revisión en las expectativas condicionada a la llegada de un determinado tipo de orden.

Por su parte, el precio de transacción sin contemplar el redondeo es el resultado de ajustar el valor cierto del activo por el beneficio que el oferente de liquidez ha de percibir por sus servicios como suministrador de immediatez (componente transitorio), resultado de comprar a un precio más bajo que el valor cierto y vender a un precio más alto. El precio observado de la transacción,  $P_t^o$ , no coincidirá con el precio en  $t$  de la transacción, anteriormente descrito, a causa del redondeo.

Respecto del componente de selección adversa y el componente transitorio de la horquilla, se plantean como una relación lineal y positiva del volumen negociado en la  $t$ -ésima transacción,  $V_t$ , como se indica en las expresiones [3] y [4].

$$Z_t = z_0 + z_1 V_t \quad [3]$$

$$C_t = c_0 + c_1 V_t \quad [4]$$

También se asume que los cambios en el valor cierto del activo consecuencia de la información pública generada entre transacciones,  $E_t$ , siguen distribuciones normales, independientes e idénticamente distribuidas. El valor esperado y la varianza de dicha variable dependerán del tiempo transcurrido entre transacciones.

<sup>3</sup> Se asume que si lo que llega es una orden de compra la revisión será al alza y al contrario para la llegada de una orden de venta.

A partir de las anteriores expresiones, y recogiendo el efecto del redondeo con la variable  $r_t$ , el cambio en el precio de transacción observado vendría dado por la expresión [5].

$$\Delta P_t^o = P_t^o - P_{t-1}^o = M_t + C_t Q_t + r_t - (M_{t-1} + C_{t-1} Q_{t-1} + r_{t-1}) \quad [5]$$

Teniendo en cuenta el valor de  $M_t$  en función de  $M_{t-1}$ , expresión [1], y considerando la forma funcional impuesta al componente transitorio y al de selección adversa, expresiones [3] y [4], la variación de los precios observado de transacción quedaría reflejada en la expresión [6].

$$\Delta P_t^o = c_0(Q_t - Q_{t-1}) + c_1(V_t Q_t - V_{t-1} Q_{t-1}) + z_0 Q_t + z_1 Q_t V_t + E_t + r_t - r_{t-1} \quad [6]$$

En el caso de conocer la dirección de la orden,  $Q_t$ , a partir de esta modelización, y calculando únicamente la variación de los precios observados entre transacciones, se pueden estimar los distintos componentes de la horquilla por mínimos cuadrados ordinarios. Ahora bien, aunque los parámetros estarían bien definidos, la estimación sería ineficiente debido a los errores de redondeo,  $r_t$ , y al hecho de que la varianza de la innovación en los precios,  $E_t$ , no es constante sino que depende del tiempo entre transacciones.

La propuesta de los autores para la estimación del modelo descrito es utilizar el procedimiento de maximoverosimilitud, ya que, en primer lugar, permite estimar la horquilla efectiva desde las series de precios sin clasificarlas como mejores precios de compra o de venta, es decir, sin determinar  $Q_t$ . En segundo lugar, y al igual que los modelos basados en las covarianzas, permite identificar el componente transitorio de la horquilla desde la reversión de los precios. Además, este método de estimación permite corregir el sesgo que produce en la horquilla estimada el no considerar el redondeo.<sup>4</sup>

Bajo ciertas hipótesis, el modelo inicialmente planteado, expresión [5], se simplifica notablemente facilitando su utilización y estimación empírica. En concreto, (i) si se asume que la media y la varianza de la variable que recoge la información pública generada entre transacciones es constante, y, por tanto, independiente de  $T$ ; (ii) que no existe redondeo y, en consecuencia, que la evolución de los precios de transacción es continua; (iii) que, como sugiere la teoría de información asimétrica, el componente de selección adversa es prácticamente despreciable para transacciones muy pequeñas; y (iv) por último, que el componente transitorio de la horquilla es constante, el modelo se simplifica considerablemente, llegando a la expresión [7].

$$\Delta P_t = c_0(Q_t - Q_{t-1}) + z_1 Q_t V_t + e_t \quad [7]$$

El modelo se podría relacionar con aquellos otros que consideran y ajustan por el coste de inventario, donde el volumen de la transacción aparece desfasada un periodo, en lugar de variable contemporánea. Si volvemos a incluir en la estimación del mismo el valor de  $c_j$ , tal y como se recoge en la expresión [8].

$$\Delta P_t = c_0(Q_t - Q_{t-1}) + b \cdot Q_{t-1} V_{t-1} + z Q_t V_t + e_t \quad [8]$$

En la expresión [8]  $b = -c_j$  y  $z = z_j + c_j$ . Por tanto, la estimación de  $b$  debería reflejar el ajuste debido a los costes de inventario.

<sup>4</sup> El hecho de que los precios sean discretos también puede causar correlación serial negativa en la serie de cambios de precios.

Los contrastes empíricos efectuados por los autores reflejan que estas últimas especificaciones más complejas no presentan ventajas frente a la más sencilla recogida en [7]. Los resultados empíricos obtenidos a partir de su modelo muestran una clara evidencia a favor de la existencia de un componente de selección adversa (ya que resulta ser positivo y significativo para la mayoría de los activos analizados), que además es función del tamaño de la orden como así determinan los modelos teóricos basados en la existencia de asimetrías informativas entre los agentes.

El modelo *Lin et al. (1995)* plantea, a diferencia del anterior, la estimación de tres componentes en la horquilla cotizada. Dos de estos componentes serían el de selección adversa, presente en todos los modelos recogidos en este trabajo y el componente de procesamiento de órdenes, cuya función sería la remuneración del oferente de liquidez. El tercer componente que se incorpora y que es característico de este modelo equivale a los costes de inventario sugerido por la teoría de microestructura<sup>5</sup> y refleja el coste relacionado con la persistencia del tipo de orden, esto es, con la probabilidad de órdenes consecutivas del mismo signo.

El modelo *Lin et al. (1995)* recoge, frente al modelo de *Glosten y Harris (1988)*, la posibilidad de que todas las órdenes no sean ejecutadas a los mejores precios de compra o venta. Por ello se distingue entre horquilla cotizada, efectiva y realizada. La diferencia entre la horquilla cotizada y la efectiva vendrá dada por las órdenes que no se ejecuten a los mejores precios cotizados, mientras que la horquilla realizada cuantificará el beneficio o ganancia esperada por el oferente de liquidez.

Si denotamos con  $\delta$  a la probabilidad de que una orden de compra/venta sea seguida por otra del mismo signo, la probabilidad de un cambio en el tipo de orden será  $(1 - \delta)$ . Si la orden ejecutada en  $t$  fue una orden de venta y, por tanto, se efectuó al mejor precio de compra el beneficio esperado por el oferente de liquidez,  $B_t$ , vendrá recogido a través de la expresión [9].

$$\delta(B_{t+1} - B_t) + (1 - \delta)(A_{t+1} - B_t) = E(P_{t+1}) - P_t \quad [9]$$

donde:  $A_{t+1}$  recoge el mejor precio de venta en  $t+1$ ,  $E(P_{t+1})$  es el valor esperado del activo en  $t+1$  condicionado a la negociación en  $t$ , y  $P_t$  es el precio de transacción en  $t$ , es decir,  $B_t$ .<sup>6</sup>

La horquilla efectiva,  $z_t$ , se calcula a través de la diferencia entre el precio de transacción y el valor medio de la horquilla cotizada,  $M_t$ . El signo de  $z_t$  dependerá del tipo de orden ejecutada. Si es una orden de venta, lo que supone una compra para el oferente, será negativo y positivo si la orden ejecutada es de compra, es decir, una venta del oferente de liquidez. En el caso de que todas las órdenes se ejecuten a los precios cotizados, el valor de  $z_t$  sería exactamente la mitad de la horquilla y no habría diferencias entre la horquilla cotizada y la horquilla efectiva.

Al igual que en otros modelos, para recoger la información transmitida por la transacción en  $t$  se ajustan los precios cotizados en  $t+1$ , siendo el ajuste una proporción  $\lambda$  de la horquilla efectiva. Se asume que la ejecución de una orden de venta provocará una disminución de los precios y al contrario si la orden ejecutada es de compra.

$$\begin{aligned} B_{t+1} &= B_t + \lambda z_t \\ A_{t+1} &= A_t + \lambda z_t \end{aligned} \quad [10]$$

<sup>5</sup> Este componente trata de compensar al oferente de liquidez por alejarse de su cartera óptima al atender las órdenes de compra y venta.

<sup>6</sup> Se está suponiendo que la transacción se efectuó a los precios cotizados.

A partir del ajuste descrito, y que se recoge en la expresión [10], los autores derivan una relación entre horquilla efectiva y horquilla realizada, tal y como se expresa en [11], que refleja, en primer lugar, la relación entre ambas horquillas si la transacción en  $t$  se efectuó al mejor precio de venta,  $A_t$ , y, en segundo lugar, la relación que se deriva condicionada a que el precio al que se ejecutó la orden en  $t$  fue el mejor precio para la compra,  $B_t$ .

$$\begin{aligned} E(P_{t+1}) - P_t &= (1 - \lambda - \vartheta)z_t = \gamma z_t \\ E(P_{t+1}) - P_t &= -(1 - \lambda - \vartheta)z_t = -\gamma z_t \end{aligned} \quad [11]$$

El beneficio esperado por el oferente de liquidez se identifica con una proporción de la horquilla efectiva que coincide con el coste de procesamiento de órdenes,  $\gamma = 1 - \lambda - \vartheta$ , donde  $\lambda$  representa, como ya hemos indicado, el ajuste de la cotización ante la transacción ocurrida y se corresponde con el componente de selección adversa, y el parámetro  $\vartheta$ , que es función de la probabilidad de continuación del tipo de orden ( $\vartheta = 2\delta - 1$ ), identifica la persistencia del flujo de órdenes.

Para la estimación de los tres componentes ya señalados que el modelo contempla en la horquilla se proponen las ecuaciones recogidas en las expresiones [12], [13] y [14].

$$M_{t+1} - M_t = \lambda z_t + e_{t+1} \quad [12]$$

$$z_{t+1} = z_t + \eta_{t+1} \quad [13]$$

$$P_{t+1} - P_t = -\gamma z_t + u_{t+1} \quad [14]$$

donde:  $e_t$ ,  $\eta_t$ ,  $u_t$  son perturbaciones aleatorias normales y no correlacionadas, luego podemos estimar los parámetros de interés por mínimos cuadrados ordinarios.

Los resultados obtenidos mediante la contrastación empírica del modelo muestran un componente de selección adversa significativo y positivo para cualquier submuestra e intervalo del día analizado.

El objetivo de *Huang y Stoll (1997)* es construir un modelo para estimar los tres componentes señalados por los modelos teóricos para la horquilla cotizada en series temporales que, de algún modo, recoja los modelos existentes.

Las principales ventajas de este modelo son la simplicidad, ya que puede ser implementado con un procedimiento de regresión en un paso, y la flexibilidad de su estructura que le permite adaptarse a determinados análisis como, por ejemplo, la relación de los distintos componentes con el tamaño negociado.

La secuencia de los acontecimientos en el momento  $t$  seguida en el modelo es la siguiente:

- Se determina el valor fundamental del activo en  $t$ , al que llamaremos  $V_t$ .
- Se fijan las cotizaciones,  $A_t$  y  $B_t$ , a partir de las cuales se calcula el punto medio de la horquilla,  $M_t$ .
- Se efectúa la transacción en  $t$  y se fija el precio de transacción,  $P_t$ .

El valor fundamental del activo en  $t$  vendrá dado por el valor en  $t-1$  ajustado por la información transmitida en la última transacción y por la llegada de información pública, tal y como se indica en la expresión [15].

$$V_t = V_{t-1} + \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t \quad [15]$$

donde:  $S$  es la horquilla cotizada que se supone constante en el tiempo,  $\alpha$  es la proporción de la misma atribuida a la selección adversa y  $\varepsilon_t$  es una perturbación aleatoria. Por otra parte,  $Q_t$  es el indicador de la dirección de la orden que, al igual que en casos anteriores, toma valor  $+1$  cuando la orden ejecutada fue de compra,  $-1$  si la orden ejecutada es de venta y  $0$  si no puede clasificarse. Una transacción se identifica como una orden de compra si se realiza por encima del punto medio de la horquilla y con una orden de venta si ejecuta por debajo del mismo.

El cambio en el valor fundamental del activo,  $\Delta V_t$ , se descompone en dos partes: la primera de ellas, recogida en el primer sumando de la expresión [16], se debe a la información privada revelada por la transacción en  $t-1$ , y la segunda parte, debida a la información pública, es capturada por  $\varepsilon_t$ .

$$\Delta V_t = \alpha \frac{S}{2} Q_{t-1} + \varepsilon_t \quad [16]$$

Desde el punto de vista de los modelos que, como éste, contemplan los costes de inventario, los precios cotizados y, por tanto, el punto medio de la horquilla se ajustan por el oferente de liquidez para inducir negociaciones que equilibren su posición, esto es, incrementará los precios cotizados después de ejecutar una orden de compra y los disminuirá en caso contrario. Desde el punto de vista de estos ajustes, si  $\beta$  es la proporción de la horquilla debida a los costes de inventario, el punto medio de la horquilla en  $t$ ,  $M_t$ , se puede expresar en función del valor fundamental del activo ajustado convenientemente para provocar la transacción del signo deseado, ajuste que depende de la posición del oferente de liquidez.

$$M_t = V_t + \beta \frac{S}{2} \sum_{i=1}^{t-1} Q_i \quad [17]$$

donde:  $\sum_{i=1}^{t-1} Q_i$  es el inventario acumulado desde que abrió el mercado hasta  $t-1$ .<sup>7</sup>

Al suponer que la horquilla es constante en el tiempo, la expresión [17], que se establece para el punto medio de la horquilla, es válida para los precios cotizados de compra,  $B_t$ , y para los precios cotizados de venta,  $A_t$ . Asimismo, la variación en el punto medio de la horquilla y en los precios cotizados vendría dada por la expresión [18].

$$\Delta M_t = \Delta V_t + \beta \frac{S}{2} Q_{t-1} \quad [18]$$

Considerando que la variación en el valor fundamental del activo es la indicada en la expresión [18], llegaríamos a establecer la variación en las cotizaciones tal y como se expresa en [19].

$$\Delta M_t = (\alpha + \beta) \frac{S}{2} Q_{t-1} + \varepsilon_t \quad [19]$$

El precio de la transacción, expresión [20], vendrá dado por la suma del punto medio de la horquilla más/menos<sup>8</sup> la mitad de la misma y más un término de error,  $h_t$ , que reco-

<sup>7</sup> Se está suponiendo que hasta  $t-1$  las transacciones son de volumen unitario.

<sup>8</sup> Dependiendo de si se ejecuta una orden de compra o de venta, respectivamente



gerá la desviación de la horquilla realmente observada de la cotizada y que incluye los posibles redondeos asociados al hecho de que los precios de transacción no sean continuos.

$$P_t = M_t + \frac{S}{2} Q_t + \eta_t \quad [20]$$

En este modelo también se distingue entre horquilla negociada,  $z$ , y cotizada,  $S_t$ . La diferencia entre ambas, al igual que en el modelo anterior, se debe a las negociaciones que no se efectúan a los precios cotizados sino en el interior de la horquilla.<sup>9</sup>

A partir de las expresiones [19] y [20], y si denominamos  $\lambda = \alpha + \beta$  y  $e_t = \varepsilon_t + \Delta\eta_t$ , la variación en el precio de transacción,  $\Delta P_t$ , que representa el modelo básico de estos autores vendría dado por la expresión [21].<sup>10</sup>

$$\Delta P_t = \frac{z}{2}(Q_t - Q_{t-1}) + \lambda \frac{z}{2} Q_{t-1} + e_t \quad [21]$$

La estimación de la expresión [21], proporciona el valor de  $z/2$  y de  $\lambda(z/2)$ , y, por tanto, permite estimar la proporción de la horquilla debida a los costes de procesamiento de órdenes,  $(1-\lambda)$ , y la horquilla efectiva o negociada,  $z$ . Sin embargo, no permite separar el componente de selección adversa y el coste de inventario.

Para llegar a una versión extendida del modelo básico que permita estimar los componentes por separado,<sup>11</sup> es necesario introducir la probabilidad de que dos órdenes seguidas sean del mismo tipo,  $(1-\pi)$ , o sean diferentes,  $\pi$ . De este modo, la esperanza de que en un momento del tiempo la transacción efectuada sea de un tipo u otro condicionada al tipo de transacción del momento anterior vendría dada por la expresión [22].

$$E(Q_{t-1} | Q_{t-2}) = (1 - 2\pi)Q_{t-2} \quad [22]$$

La consideración de estas probabilidades en la secuencia de órdenes implica variaciones en las expresiones de la variación del valor fundamental, expresión [16], y del punto medio de la horquilla, expresión [19]. Ahora, dichas variaciones vendrán dadas por las expresiones [23] y [24], respectivamente.

$$\Delta V_t = \alpha \left[ \frac{z}{2} Q_{t-1} - \frac{z}{2} (1 - 2\pi) Q_{t-2} \right] + \varepsilon_t \quad [23]$$

$$\Delta M_t = (\alpha + \beta) \frac{z}{2} Q_{t-1} - \alpha \frac{z}{2} (1 - 2\pi) Q_{t-2} + \varepsilon_t \quad [24]$$

Al contrario que la variación en el valor fundamental del activo, la variación en el valor del punto medio de los precios cotizados no se altera por la introducción de las probabilidades en la secuenciación del tipo de órdenes porque el riesgo de inventario no se ve

<sup>9</sup> La posibilidad de negociaciones fuera de la horquilla no se contempla. Si éstas ocurren son ignoradas para la estimación del modelo.

<sup>10</sup> El modelo básico planteado puede generalizarse para realizar estimaciones diferentes dependiendo, por ejemplo, del momento de la transacción o del tamaño de la misma.

<sup>11</sup> La corrección en las cotizaciones basadas en los modelos de costes de inventario induce correlación negativa en la secuencia de tipos de orden, ya que después de efectuar una compra el oferente de liquidez baja sus cotizaciones con el fin de que la siguiente transacción sea una venta. Por ello, la extensión del modelo básico se basa en esta autocorrelación negativa en el flujo de órdenes.

afectado por la esperanza de la realización de una determinada negociación sino por la negociación realmente efectuada.

La variación en los precios de transacción bajo este nuevo contexto sería la que se refleja en la ecuación [25], que es análoga a la del modelo básico, expresión [21], pero a diferencia de ésta, permite estimar los tres componentes de la horquilla además de la horquilla negociada o efectiva. Todos estos parámetros, junto con la probabilidad de que exista un cambio en la secuencia de órdenes ejecutadas, se obtienen estimando simultáneamente las expresiones [22] y [25].

$$\Delta P_t = \frac{z}{2}(Q_t - Q_{t-1}) + \alpha \frac{z}{2} Q_{t-1} + \beta \frac{z}{2} Q_{t-1} - \alpha \frac{z}{2} (1 - 2\pi) Q_{t-2} + e_t \quad [25]$$

La variación en el punto medio de la horquilla también permite estimar los componentes de la misma y la probabilidad de cambio en el tipo de orden. Para ello se parte de los valores de la horquilla cotizada, sin diferenciar entre éstas y la horquilla efectiva o negociada. En este caso, la expresión a estimar es la que aparece en [26].

$$\Delta M_t = \alpha \frac{S_{t-1}}{2} Q_{t-1} + \beta \frac{S_{t-1}}{2} Q_{t-1} - \alpha (1 - 2\pi) \frac{S_{t-1}}{2} Q_{t-2} + e_t \quad [26]$$

Otra propuesta de los autores para estimar los componentes de la horquilla por separado se basa en que el problema de selección adversa es específico de un activo mientras que el riesgo de inventario no dependerá de la negociación sobre un solo activo sino de la negociación del resto de activos que componen la cartera del oferente de liquidez, ya que lo que se pretende es mantener el equilibrio global de la cartera. Así, la compra de un determinado activo por parte del oferente de liquidez puede provocar no solamente la bajada de los precios cotizados de ese activo, sino también de todos aquellos correlacionados con él, o incluso no provocar variaciones a la baja si sobre otros activos hay exceso de compra.

La consideración del efecto de la negociación sobre toda la cartera del oferente de liquidez, y no solamente la negociación de un activo individual, modifica la ecuación que relaciona el punto medio de la horquilla con el valor de equilibrio del activo, expresión [19], ya que como indicador de la negociación no se ha de tomar la serie de transacciones del activo de forma aislada sino el agregado de las transacciones de los activos que conforman la cartera, como se refleja en la expresión [27].

$$M_t^k = V_t^k + \beta^k \frac{z^k}{2} \sum_{i=1}^{t-1} Q_i^* \quad [27]$$

donde: el superíndice  $k$  hace referencia al activo considerado y  $Q_i^*$  es la variable indicadora del flujo de órdenes, la cual sigue tomando valores  $+1$ ,  $-1$  ó  $0$ , pero ahora este valor no depende de la transacción en  $t$  del activo  $j$ , sino del indicador agregado de las transacciones de los  $n$  activos a través de los cuales se trata de captar la dirección del mercado.

A partir de esta modificación, el modelo a estimar es el que aparece en la expresión [28]

$$\Delta P_t^k = \frac{z^k}{2} \Delta Q_t^k + \alpha^k \frac{z^k}{2} Q_{t-1}^k + \beta^k \frac{z^k}{2} Q_{t-1}^* + e_t^k \quad [28]$$

Los autores estiman empíricamente el modelo básico, diferenciando el tamaño de la orden, y los modelos extendidos, utilizando para ello el método generalizado de momen-

tos, obteniendo una vez más estimaciones para el componente de selección adversa significativas y positivas.

*Madhavan et al. (1997)* desarrollan y contrastan un modelo estructural de formación de precios, que incorpora los impactos debidos a la información pública y los efectos señalados por la microestructura. Tienen en cuenta los costes de transacción y los diferentes componentes de la horquilla, la dependencia del flujo de órdenes, la posible ejecución de negociaciones en el interior de la horquilla, la existencia de redondeo y la presencia de autocorrelación en los rendimientos del valor fundamental del activo.

Como en los otros modelos analizados, se asume que el cambio en las creencias sobre el valor fundamental del activo de los agentes entre  $t-1$  y  $t$  tiene su origen en la información pública anunciada<sup>12</sup> y en el flujo de órdenes, el cual proporciona una señal ruidosa acerca del valor futuro del activo.

Asimismo, se supone que si el oferente de liquidez cree que algunos agentes pueden poseer información privada acerca del valor fundamental del activo, tras una orden de compra se producirá una revisión al alza de dicho valor fundamental y tras una orden de venta una revisión a la baja. De esta manera, la revisión en el valor fundamental del activo está positivamente correlacionada con la innovación en el flujo de órdenes. Formalmente,<sup>13</sup> dicha revisión vendrá dada por la expresión [29], siendo el término entre paréntesis la sorpresa en el flujo de órdenes.

$$\theta[Q_t - E(Q_t | Q_{t-2})] \quad [29]$$

donde:  $Q_t$  indica, al igual que en el resto de modelos el signo de la orden, el cual tomará valor  $+1$  si la transacción la originó una orden de compra,  $-1$  si la originó una orden de venta y  $0$  en caso de que no pueda calificarse porque se ejecutó en el interior de horquilla, y  $\theta$  representa el componente de selección adversa de los costes de transacción, o lo que es lo mismo, el impacto permanente de la innovación del flujo de órdenes.<sup>14</sup>

Para describir el comportamiento de los precios de transacción,  $P_t$ , y de las cotizaciones,  $A_t$  y  $B_t$ , se supone inicialmente un tamaño fijo para la orden, lo que permite derivar en un modelo más sencillo.<sup>15</sup>

Si denotamos con  $V_t$  al valor fundamental esperado del activo después de la negociación en  $t$  condicionado a la información pública aparecida y al flujo de órdenes, su valor vendrá dado por la ecuación [30].

$$V_t = V_{t-1} + \theta(Q_t - E(Q_t | Q_{t-1})) + \varepsilon_t \quad [30]$$

Los precios para la compra,  $B_t$ , y para la venta,  $A_t$ , que se fijan antes de la negociación recogerán, como se observa en la expresión [31], además de la revisión del valor

<sup>12</sup> El cambio en las creencias respecto al valor fundamental del activo debido a la nueva información pública anunciada entre  $t-1$  y  $t$  viene recogido por una variable aleatoria,  $e_t$ , independiente e idénticamente distribuida con media cero y varianza constante.

<sup>13</sup> A diferencia del modelo de Huang y Stoll (1997) la revisión en el valor fundamental del activo se realiza no en función del valor de  $Q_t$  sino del valor no esperado de dicha variable.

<sup>14</sup> La revisión del valor fundamental del activo a partir de este parámetro también se contempla en el modelo de Hasbrouck (1991) al que haremos referencia a continuación.

<sup>15</sup> A favor de este supuesto argumentan que el agente que trate de ejecutar una orden grande la fraccionará para introducirla en el mercado, o si no este tipo de órdenes se ejecutan fuera del mercado para evitar el impacto excesivo de las mismas y para encontrar contrapartida más rápidamente.

fundamental del activo ya indicada, un componente que compense al oferente de liquidez por sus servicios,  $f$ . Este componente recoge conjuntamente el coste de procesamiento de órdenes y el coste de inventario imputado por acción.

$$\begin{aligned} A_t &= V_{t-1} + \theta[1 - E\langle Q_t | Q_{t-1} \rangle] + \phi + \varepsilon_t \\ B_t &= V_{t-1} + \theta[1 - E\langle Q_t | Q_{t-1} \rangle] + \phi + \varepsilon_t \end{aligned} \quad [31]$$

Los autores asumen que todas las órdenes se ejecutan o al mejor precio de compra o al mejor precio de venta o al precio medio de la horquilla.

El precio de transacción,  $P_t$ , vendrá dado por la expresión [32], donde  $\xi_t$  es una variable independiente e idénticamente distribuida con media cero que captura posibles errores debidos al redondeo y posibles variaciones de los rendimientos a lo largo del tiempo.

$$P_t = V_t + \phi Q_t + \xi_t \quad [32]$$

Dado que se supone que el oferente de liquidez actúa bajo la hipótesis de expectativas racionales, establece sus cotizaciones en función de la negociación venidera de tal manera que el precio de la transacción en  $t$  ya incorpora el efecto de la dirección en dicha transacción. Por eso se establece en función de  $V_t$ , que es el valor del activo condicionado a que la transacción en  $t$ ,  $Q_t$ , ha sido observada.

Sustituyendo el valor fundamental del activo en  $t$  por su valor en función de  $t-1$  la expresión quedaría como aparece en [33].

$$P_t = V_{t-1} + \theta(Q_t - E\langle Q_t | Q_{t-1} \rangle) + \phi Q_t + \varepsilon_t + \xi_t \quad [33]$$

Para poder estimar la expresión [33] se debe sustituir la variable no observable, el valor fundamental del activo, por variables que puedan ser observadas a través de los datos. Para ello, se debe describir el comportamiento temporal del flujo de órdenes, variable para la cual se asume un proceso general de Markov.

La probabilidad de que dos órdenes consecutivas sean del mismo signo,  $g$ , es decir, que la negociación sea iniciada por una orden de compra/venta condicionada a que la anterior fuera del mismo signo, es la que se indica en [34]. Si se parte de la hipótesis de que las órdenes de un tamaño relativamente grande se introducen en el mercado fraccionadas en órdenes más pequeñas para facilitar su ejecución,  $\gamma > 1/2$ , entonces, la probabilidad de encontrarse con dos órdenes del mismo signo es mayor que la probabilidad de cambio.

$$\gamma = \Pr\langle Q_t = Q_{t-1} | Q_{t-1} \neq 0 \rangle \quad [34]$$

Si llamamos  $l$  a la probabilidad no condicionada de que la negociación se efectúe en el interior de la horquilla, es decir, la probabilidad de que  $Q_t$  sea igual a cero, el coeficiente de autocorrelación de primer orden de la variable indicadora de la dirección de la orden,  $r$ , viene expresado en [35].

$$\rho = 2\gamma - (1 - \lambda) \quad [35]$$

Cuando no hay posibilidad de transacción dentro de la horquilla,  $\lambda$  es igual a cero y  $\rho = 2\gamma$ . Si, además, las órdenes son independientes,  $\gamma = 1/2$ , el coeficiente de autocorrelación será nulo.

La esperanza condicional de la dirección de la negociación en el momento  $t$  condicionada al valor que tomó en  $t-1$  se establece en [36].

$$\begin{aligned} Q_{t-1} = 0 &\Rightarrow E\langle Q_t | Q_{t-1} = 0 \rangle = 0 \\ Q_{t-1} = 1 &\Rightarrow E\langle Q_t | Q_{t-1} = 1 \rangle = \gamma - (1 - \gamma - \lambda) = 2\gamma - (1 - \lambda) = \rho \\ Q_{t-1} = -1 &\Rightarrow E\langle Q_t | Q_{t-1} = -1 \rangle = -\gamma + (1 - \gamma - \lambda) = -2\gamma + (1 - \lambda) = -\rho \end{aligned} \quad [36]$$

De las expresiones contenidas en [36], podemos deducir que la esperanza de la dirección de la orden en el momento  $t$  condicionada a cualquier valor de la dirección de la negociación en  $t-1$  viene dada por [37].

$$E\langle Q_t | Q_{t-1} \rangle = \rho Q_{t-1} \quad [37]$$

Partiendo de la probabilidad condicional calculada en [34] y de la relación entre el valor fundamental del activo y el precio de transacción reflejada en [32], el modelo a estimar puede expresarse como aparece en [38].

$$P_t - P_{t-1} = (\phi + \theta)Q_t - (\phi + \rho\theta)Q_{t-1} + \varepsilon_t + \xi_t + \xi_{t-1} \quad [38]$$

A partir del modelo desarrollado se muestra cómo estimar los costes de transacción medidos por la horquilla cotizada o la efectiva a partir de los parámetros  $\theta$  y  $\phi$ , que como ya se indicó, recogen, respectivamente, los efectos de la información asimétrica y los costes de procesamiento de órdenes e inventario.

La horquilla cotizada para el momento  $t$ , ex-ante es una variable aleatoria cuya media es  $2(\theta + \phi)$ . La estimación de dicha media vendría dada por la expresión calculada con los valores estimados de los parámetros del modelo mediante el método generalizado de momentos.

La otra medida del coste de transacción la daría la horquilla realizada, que se define como la esperanza del cambio en el precio entre una compra nocional en  $t$  y una venta nocional en  $t+k$ . Para definir esta horquilla se asume que el tiempo transcurrido entre  $t$  y  $t+k$  es suficiente como para suponer que han transcurrido varias transacciones y que al final del intervalo existe la misma probabilidad de que se presente una compra o una venta con el fin de ignorar la autocorrelación del flujo de órdenes. La horquilla realizada,  $s$ , vendría dada por la expresión [39].

$$s = (1 - \lambda)(2\phi + \theta) \quad [39]$$

A partir del modelo desarrollado también se establece la volatilidad de los precios de transacción y la autocovarianza entre los cambios de dichos precios. La volatilidad del precio, medida por la varianza de los cambios en dicha variable, puede ser utilizada para medir el impacto de las transacciones en el precio. Bajo el modelo planteado se verifica que la varianza de los cambios en los precios es igual al valor indicado en la expresión [40].

$$Var(\Delta P_t) = \sigma_\varepsilon^2 + 2\sigma_\xi^2 + (1 - \lambda)[(\theta + \phi)^2 + (\theta\rho + \phi)^2 - 2(\theta + \phi)(\theta\rho + \phi)\rho] \quad [40]$$

La expresión [40] permite dividir la volatilidad de los cambios en los precios en dos componentes: uno de ellos ligado a la información pública y el otro a las transacciones. Respecto al primero de estos componentes,  $\sigma_\varepsilon^2$ , recoge la volatilidad debida a las nuevas noticias. El segundo componente recoge: (i) la parte de la volatilidad que se debe a la no continuidad del precio de transacción, es decir, al redondeo,  $2\sigma_\xi^2$ ; (ii) la parte debida a la

información privada transmitida por la negociación,  $(1 - \lambda)(1 - \rho^2)\theta^2$ ; (iii) la parte debida a los costes de inventario y procesamiento de órdenes,  $2(1 - \lambda)(1 - \rho)\phi^2$ ; y (iv) un término al que los autores denominan de interacción y donde aparecen ambos parámetros,  $2\phi\theta(1 - \lambda)(1 - \rho^2)$ .

Esta descomposición muestra que la volatilidad de los precios mantiene, por una parte, una relación directa con los componentes de los costes e inversa con la autocorrelación del flujo de órdenes y, por otra, con la probabilidad de que las órdenes sean ejecutadas en el interior de la horquilla.

A partir de la varianza de los cambios en los precios se puede calcular la proporción de volatilidad de los precios que es debida a la aparición de nuevas noticias, y la proporción que es debida a las fricciones del mercado (redondeo, costes de transacción,...). La fracción de la varianza atribuible a las fricciones de la negociación viene dada en la expresión [41].

$$\pi = \frac{2\sigma_{\xi}^2 + (1 - \lambda)[(\theta + \phi)^2 + (\theta \cdot \rho + \phi)^2 - 2(\theta + \phi)(\theta \cdot \rho + \phi)\rho]}{\sigma_{\varepsilon}^2 + 2\sigma_{\xi}^2 + (1 - \lambda)[(\theta + \phi)^2 + (\theta \cdot \rho + \phi)^2 - 2(\theta + \phi)(\theta \cdot \rho + \phi)\rho]} \quad [41]$$

De forma análoga, se puede calcular la proporción de la volatilidad debida al efecto de la información asimétrica o de cualquiera de los otros tres componentes señalados arriba simplemente sustituyendo en el numerador de la expresión [41] la parte de la varianza que le corresponde a cada elemento.

El modelo recoge la autocorrelación existente entre la variación de los precios de transacción y la variación de las cotizaciones que se debe a las fricciones del mercado. En concreto, la autocorrelación de los rendimientos de transacción es negativa si existe redondeo y costes asociados al procesamiento de órdenes y al riesgo de inventario. El valor absoluto de la autocovarianza crece con las fricciones del mercado y cuanto mayor es la presencia de información asimétrica, y disminuye cuanto mayor es la probabilidad de la existencia de órdenes que se ejecuten en el interior de la horquilla.

Por otra parte, la autocovarianza de las variaciones de las cotizaciones de venta, o de compra (el modelo es simétrico y, por tanto, las conclusiones derivadas para un caso son válidas para el otro), y a pesar de que le oferente de liquidez actúa bajo una estructura de expectativas racionales, puede ser distinta de cero. Es más, solamente será nula en el caso de que no existan fricciones debidas al redondeo y que no existan costes asociados a la información asimétrica, o en el caso de que la autocorrelación del flujo de órdenes sea nula o todas las órdenes se ejecuten en el interior de la horquilla.

Los resultados obtenidos por Madhavan *et al.* (1997) en la estimación del modelo muestran que, aunque las estimaciones difieren de un título a otro, todas ellas son significativamente distintas de cero. Asimismo, comprueban que la estimación de los costes ligados a la información asimétrica decrece en el curso de la jornada, mientras que los costes de transacción incrementan durante el periodo de negociación. Por otra parte la autocorrelación en el flujo de órdenes es estable.

Hasbrouck (1991) propone una modelización de las interacciones sugeridas por los modelos teóricos de microestructura entre las negociaciones y la revisión de las cotizaciones mediante vectores autorregresivos, con la finalidad de recoger de forma más idónea la dinámica por la cual las negociaciones y las cotizaciones interactúan.

El mismo autor, en un trabajo anterior [Hasbrouck (1988)], concluye que los efectos de la información y del coste de inventario que producen cambios en las cotizaciones a

partir de las negociaciones podían ser parcialmente resueltos atendiendo a la persistencia de su impacto sobre el precio del activo. En este último trabajo reconoce otros efectos producidos por determinadas imperfecciones en el mercado: redondeo, fragmentación de las órdenes, presión de los precios, que, como el coste de inventario, tienen un efecto transitorio en la determinación de las cotizaciones.

Al igual que en el trabajo de Madhavan *et al.* (1997), en el presente modelo se contempla el hecho de que la información inferida por la negociación proviene exclusivamente de la parte no anticipada de la misma, y define formalmente el impacto de la información de una negociación como el efecto persistente en el precio del activo resultante del componente no esperado de la transacción. Dicho efecto en el precio ha de ser estudiado desde el impacto inmediato ya que más tarde puede estar contaminado por los efectos transitorios producidos por el control de inventario y otras imperfecciones del mercado.

En cuanto a la secuencia de acontecimientos, supone que las cotizaciones de compra y venta,  $A_t$  y  $B_t$ , se fijan después de la negociación en  $t$  y, por tanto, recogen la información transmitida por la negociación  $t$ -ésima.

La negociación es caracterizada por una variable,  $x_t$ , que recoge conjuntamente el signo de la orden y el volumen. A partir de esta variable el oferente de liquidez fija las cotizaciones que se suponen simétricas alrededor del valor esperado del activo condicionado a toda la información pública, incluida la transmitida por la negociación. Por ello, como valor fundamental del activo se toma el precio medio de la horquilla.

La hipótesis de simetrías de las cotizaciones respecto al valor fundamental del activo vendría reflejada formalmente en la expresión [42], donde  $t$  es un momento final futuro en el tiempo y  $z_t$  es el valor fundamental del activo en ese momento final condicionado a toda la información pública disponible en  $t, f_t$ .

$$E\langle (A_t + B_t)/2 - \zeta_t \mid \phi_t \rangle = (A_t + B_t)/2 - E\langle \zeta_t \mid \phi_t \rangle = 0 \quad [42]$$

Bajo esta hipótesis, la información inferida por la negociación en  $t$ ,  $x_t$ , puede contemplarse como la subsiguiente revisión en el valor medio de las cotizaciones,  $r_t$ .

$$r_t = [(A_t + B_t)/2] - [(A_{t-1} + B_{t-1})/2] \quad [43]$$

Los costes de transacción, los cuales se consideran fijos y simétricos para la compra y para la venta, se acomodan a este punto.

Se asume que, si existe información pública distinta a la desprendida de la transacción en  $t$ , ésta llega después de la  $t$ -ésima transacción y antes de la revisión de las cotizaciones. Si esto es así, la revisión de las cotizaciones refleja información pública y privada, lo que impide que se pueda conocer el impacto en el precio de una negociación en particular. Por ello se supone que la revisión de las cotizaciones es una función estable de la negociación y, además, que la dependencia es lineal, tal y como se recoge en [44].

$$r_t = bx_t + v_{1t} \quad [44]$$

donde:  $v_{1t}$  refleja la información pública, mientras que  $b$  mide el impacto que la negociación tiene en la revisión de la cotización.

La especificación [44] considera solamente el impacto contemporáneo que tiene la negociación sobre la determinación de las cotizaciones. Sin embargo, para construir un modelo más robusto, plantea la necesidad de reconocer que existen imperfecciones pro-

pias del funcionamiento del mercado que provocan ajustes de información retardados, es decir, que la revisión en la cotización dependerá no sólo de la negociación contemporánea sino de negociaciones anteriores, por ello establece para la revisión de la cotización una estructura más flexible donde aparecen valores retardados de las variables, como se observa en [45].

$$r_t = a_1 r_{t-1} + a_2 r_{t-2} + \dots + b_0 x_t + b_1 x_{t-1} + \dots + v_{1t} \quad [45]$$

La correlación serial en la revisión de las cotizaciones es incompatible con la hipótesis de cotizaciones simétricas expuesta antes, ya que el modelo descrito sobre la base de variables retardadas permite desviaciones de los precios eficientes respecto al valor fundamental del activo. Estas desviaciones serán transitorias, por lo que con el paso del tiempo los agentes, que se comportan de forma racional, esperan que las cotizaciones en media reviertan al valor cierto del activo. Siguiendo esta idea, la hipótesis anterior sobre la simetría de las cotizaciones respecto al valor fundamental del activo condicionado a la información pública en  $t$  se sustituye por una versión más débil donde las simetrías se mantienen para valores esperados en algún momento futuro,  $f$ , tal y como se indica en [46].

$$f \rightarrow \tau, \quad E\langle (A_f + B_f)/2 - \zeta_\tau | \phi_t \rangle \rightarrow 0 \quad [46]$$

Como ya se ha dicho, las mismas imperfecciones en el funcionamiento del mercado, que provocan retrasos en los ajustes de las cotizaciones a la información desprendida de la negociación, provocan que parte de dicha negociación pueda ser anticipada. Por tanto, si en la negociación se transmite información privada, ésta debe residir en aquella parte no esperada. Para recoger este hecho el modelo se completa con la modelización de la variable indicadora de la negociación sobre la base de una serie de variables retardadas, expresión [47].

$$x_t = c_1 r_{t-1} + c_2 r_{t-2} + \dots + d_1 x_{t-1} + d_2 x_{t-2} + \dots + v_{2t} \quad [47]$$

La ecuación [47] refleja que los agentes, basándose en la historia de precios y negociaciones, construyen una determinada expectativa para la  $t$ -ésima transacción. Si existe nueva información contenida en la variable  $x_t$ , ésta debe residir en el componente no anticipado de la negociación  $v_{2t}$ , ya que el resto es totalmente conocido. Este hecho no implica que la innovación sea una función determinista de la nueva información, ya que la presencia de agentes no informados (agentes de liquidez) introduce un componente de ruido en  $u_{2t}$  que está incorrelacionado con la información privada.

Las ecuaciones [45] y [47] forman un modelo autorregresivo bivalente, en el cual se cumple que las esperanzas de las perturbaciones son nulas y que están conjunta y serialmente incorrelacionadas, tal y como se recoge en [48].

$$E(v_{1t}) = E(v_{2t}) = 0 \quad [48]$$

$$E(v_{1t}, v_{1s}) = E(v_{2t}, v_{2s}) = E(v_{1t}, v_{2s}) = 0 \quad f \neq t$$

En la primera ecuación del modelo planteado, para explicar la revisión de la cotización, se incluye el valor contemporáneo de la negociación  $x_t$ . Esto refleja el hecho de que la negociación y la revisión en la cotización no se determinan simultáneamente, y, estadísticamente, supone que la covarianza entre las perturbaciones aleatorias de ambas ecuaciones en un momento  $t$  es cero,  $E(v_{1t}, v_{2t})$ .



Como hemos indicado, la innovación de la ecuación que modeliza la negociación no es una función determinista de la información privada. Para valorar el contenido informativo de la misma hay que suponer que en un momento  $t = 0$  el modelo planteado es estable, lo que quiere decir que los valores retardados de  $r_t$  y  $x_t$  son cero y que la cotización en el momento anterior, es decir, en  $t = -1$ , es tal que su valor esperado medio es igual al valor esperado no condicional del valor fundamental del activo en  $t$ , expresión [49].

$$E[(A_{-1} + B_{-1})/2] = E[\xi_t] \quad [49]$$

A partir de aquí, la información revelada por la parte no esperada de la transacción y la medida de información asimétrica propuesta es la esperanza de la revisión de las cotizaciones acumuladas en  $m$  pasos, medida que se refleja en [50], ya que esta esperanza converge a la revisión del precio eficiente tal y como se indica en la expresión [51].

$$\alpha_m(v_{20}) = \sum_{t=0}^m E\langle r_t | v_{20} \rangle \quad [50]$$

$$\alpha_m(v_{20}) = \sum_{t=0}^m E\langle (A_t + B_t)/2 - (A_{t-1} + B_{t-1})/2 | v_{20} \rangle \rightarrow E\langle \zeta_\tau | v_{20} \rangle - E[\zeta_\tau] \quad [51]$$

La contrastación empírica del modelo muestra que el impacto completo de una transacción en el precio del activo se incorpora tras un retardo relativamente prolongado y que el impacto de la negociación es una función creciente y cóncava respecto al tamaño de la orden. Además, las negociaciones que ocurren cuando las horquillas son más amplias tienen mayor impacto en la revisión de los precios. Asimismo muestran que para las empresas pequeñas el componente de información asimétrica tiene mayor importancia.

### 3. ANÁLISIS COMPARATIVO DE LOS DISTINTOS MODELOS Y APLICABILIDAD DE LOS MISMOS

Todos los modelos analizados en este trabajo fueron desarrollados partiendo del funcionamiento de mercados dirigidos por precios. Es además en estos mercados en donde los distintos autores estiman sus modelos y donde más han sido utilizados y comparados posteriormente.

La idea de la que parten todos ellos es común: la presencia de agentes operando en el mercado con información privada provoca que las transacciones tengan contenido informativo. La información contenida en el flujo de órdenes se incorpora a las cotizaciones que son revisadas tras la transacción por el creador de mercado, y, por ende, al precio de la siguiente transacción. Este mecanismo por el cual se incorpora la información a los precios del mercado es la base para medir el grado de información asimétrica a través del efecto permanente que tiene el flujo de órdenes en los precios de transacción.

Además de la similitud señalada en el punto de partida para el desarrollo de los modelos descritos, los datos necesarios para su implementación empírica son, esencialmente, los mismos. De hecho, las conclusiones obtenidas por los distintos trabajos coinciden en líneas generales, es decir, los distintos estudios realizados coinciden en la existencia de un componente de selección adversa significativo y positivo en el mercado que además es mayor en las empresas menos negociadas y que tiene mayor importancia al prin-

cipio y al final del día. A pesar de ello, las magnitudes estimadas para dicho componente con uno u otro modelo varían notablemente (de un 9,6% en el trabajo de Huang y Stoll (1997) a un 40% en el de Madhavan *et al.* (1997)). Las importantes divergencias en lo que se refiere a la magnitud de los componentes estimados plantean la importancia que puede tener la metodología utilizada a la hora de utilizar estos modelos en diferentes análisis donde el grado de información asimétrica o la formación del precio sean objeto de estudio.

Ante la falta de consenso en la magnitud de las estimaciones obtenidas en los distintos modelos, consideramos fundamental reflexionar sobre las diferencias entre los mismos y, sobre todo, si alguno de ellos se puede considerar superior al resto en su planteamiento y/o aportaciones. Las diferencias fundamentales entre los distintos modelos aquí estudiados estriban en las hipótesis iniciales, en el número de componentes que se estiman y en la técnica estadística utilizada.

Respecto al número de componentes, la comparación de los resultados obtenidos por los distintos modelos nos permite afirmar que aquellos que estiman solamente dos componentes obtienen mayor peso para el componente de selección adversa de la horquilla de precios [Madhavan *et al.* (1997)]. Este resultado es una clara evidencia de que, aplicando determinadas metodologías, se podría estar sobreestimando la negociación informada en el mercado.

Respecto a las hipótesis referidas al funcionamiento y posibilidades de negociación en el mercado que subyacen a los distintos modelos, es el modelo de Glosten y Harris (1988) el menos flexible de todos, ya que solamente recoge la posibilidad de que las órdenes puedan ejecutarse al bid y al ask y no en el interior de la horquilla por lo que este modelo estaría descartando un gran número de transacciones. Por otra parte, ninguno recoge la posibilidad de órdenes ejecutadas fuera de la horquilla que, como señalaremos más adelante, es habitual en los mercados dirigidos por órdenes como el nuestro. A la hora de modelizar el efecto en el precio de transacción a partir del flujo de órdenes, dos de los modelos descritos, Hasbrouck (1991) y Madhavan *et al.* (1997), lo hacen a través del flujo no esperado en lugar del flujo de órdenes total lo que, en principio, parece más adecuado ya que la negociación informada se identifica con la «sorpresa» de la negociación.

Todos los modelos descritos miden la selección adversa como parte de los costes de transacción, bien como un porcentaje de la horquilla cotizada bien en unidades monetarias por volumen negociado, menos el modelo de Hasbrouck (1991), el cual mide la parte de la variabilidad del valor cierto del activo que viene explicada por la parte no esperada del flujo de órdenes. El modelo de Madhavan *et al.* (1997) trata de recoger ambas medidas.

El interés que tiene el problema de la selección adversa y la importancia de cuantificar de algún modo la intensidad de la negociación informada en los mercados bursátiles ha provocado que surjan análisis cuyo único objetivo es el de comparar los resultados alcanzados con los distintos modelos descritos en este trabajo y su relación con las variables que, tradicionalmente, se han utilizado en el campo de las finanzas corporativas para aproximar situaciones donde existe mayor probabilidad de información asimétrica y, por tanto, de negociación informada (variables tales como la volatilidad, el ratio valor de mercado/valor en libros, la propiedad institucional,...). Este es el objetivo de los estudios presentados por Clarke y Shastri (2001) y Van Ness *et al.* (2001). Los resultados mostrados en ambos trabajos no permiten afirmar de forma contundente que alguno de los mo-

delos estudiados sea claramente superior a los otros, incluso muestran que el uso de la horquilla cotizada directamente observada en el mercado podría resultar igual de válido que el uso de las estimaciones de sus distintos componentes.

Por otra parte, y en lo que se refiere a la relación que los distintos estimadores mantienen con las variables indicadas por las finanzas corporativas, todos ellos presentan las relaciones esperadas con algunas de dichas variables *proxys* de la información asimétrica mientras que también presentan resultados no esperados con algunas otras. Por lo que respecta a la relación entre los distintos modelos, todos ellos proporcionan estimaciones altamente correlacionadas. En lo que se refiere a la plausibilidad de las estimaciones, son los modelos de Huang y Stoll (1997), y en menor medida el de Madhavan *et al.* (1997) los que presentan mayores problemas. Hemos de señalar que ninguno de los trabajos mencionados incluye el modelo de Hasbrouck (1991), seguramente porque la medida propuesta por el mismo difiere notablemente en concepto de las otras.

El estudio de las relaciones entre las distintas medidas utilizadas en la literatura clásica para detectar la información asimétrica en finanzas corporativas y las medidas sugeridas por los modelos descritos plantean una duda razonable sobre la posible mala especificación de los modelos y, por tanto, sobre la utilidad de los mismos frente a las *proxis* señaladas. A pesar de todo, y desde nuestro punto de vista, las medidas señaladas por los modelos empíricos de microestructura que descomponen la horquilla de precios, aunque en ningún caso guardan una relación directa con ninguna de las variables que sugieren por lógica mayor información asimétrica, tienen una clara ventaja y es que permiten estudiar el fenómeno de la presencia de negociación privada para periodos muy cortos de tiempo transacción a transacción.

En cuanto a la aplicabilidad de los distintos modelos descritos en el mercado bursátil español, como ya hemos señalado son modelos que fueron desarrollados a partir del funcionamiento de los mercados dirigidos por precios, lo cual, y aunque esté ampliamente aceptado identificar las órdenes límite del libro de órdenes como las posiciones del creador de mercado y, por tanto, como la horquilla cotizada, este aspecto no es trivial a la hora de trasladar su aplicación a mercados como el nuestro, ya que el diseño de los mismos está claramente condicionado por el funcionamiento del mercado.

Para entender la aplicabilidad que estos modelos tienen en nuestro mercado, es fundamental conocer con qué datos de frecuencia intradiaria se cuenta en el mismo. A este respecto, comentar que las tres bases de datos que actualmente se comercializan en nuestro mercado recogen:

- La primera de ella todos los cambios que se producen para el primer nivel del libro de órdenes, los precios de transacción y el volumen de la misma junto con la hora a la que se registran en el sistema.
- La segunda cuenta con información acerca de las distintas transacciones que tienen lugar en la sesión bursátil
- Y, la tercera, recoge la información de cualquier cambio que se produce en cualquiera de los cinco primeros niveles del libro de órdenes. En ésta, el campo que recoge la hora del cambio no coincide exactamente con la hora en la que fueron registrados en el sistema, siendo siempre posterior.

Dada la filosofía de los distintos modelos, a la hora de efectuar las estimaciones es fundamental secuenciar correctamente los cambios en las cotizaciones y los precios de

transacción al igual que clasificar las transacciones como de compra o de venta. En este sentido el mercado bursátil español presenta claras ventajas dado que se pueden identificar dicha secuencia sin dificultad y clasificar las órdenes trabajando con la primera base de datos mencionada: Ahora bien, eso implica que solamente se tiene información del primer nivel. En este caso el problema aparece porque nos encontraremos en numerosas ocasiones con transacciones que se ejecutan fuera de la horquilla,<sup>16</sup> supuesto éste que no se recoge en ninguno de los modelos analizados y que supondría, en contra de lo previsto, una horquilla efectiva mayor que la cotizada lo que afectaría sin duda a la estimación de los distintos componentes.

El problema sugerido con la horquilla efectiva en un mercado como el nuestro, se resolvería prácticamente en la totalidad de los casos trabajando con los precios y las profundidades de los cinco niveles y calculando la horquilla cotizada con un promedio de precios ponderados por profundidad. El uso de los cinco niveles obligaría a utilizar además otra base de datos donde se recogiese información acerca de las transacciones. Esto complica enormemente la secuenciación de las cotizaciones y las transacciones por dos motivos fundamentales: el primero es que el formato de fecha en las distintas bases de datos mencionadas cambia, y el segundo que las horas que aparecen en la base de datos de los cinco niveles del libro de órdenes no refleja, como hemos dicho, el momento exacto en el que el cambio se introduce en el sistema, sino que se recoge con un retraso respecto a dicho momento.

Otro tema importante a considerar es la agrupación de los datos originales cuando éstos provienen de la misma orden de mercado. En muchas ocasiones la misma orden de mercado es ejecutada contra más de una orden límite al mismo o distinto precio y esto supone más de un registro en las bases de datos mencionadas. Es fundamental localizar estos casos y agruparlos ya que si no es así se incrementaría la probabilidad de órdenes consecutivas del mismo signo cuando la realidad no es esa, y esto es un aspecto de considerable importancia en los modelos descritos, sobre todo en el de Lin *et al.* (1995) donde uno de los componentes de la horquilla de precios es precisamente la persistencia de órdenes y en aquellos otros en los que se supone una probabilidad del 50% de que las órdenes de un tipo sean seguidas por órdenes del mismo signo, ya que, y simplemente por el registro de datos, esta probabilidad sería más alta.

Todos los modelos descritos, excepto el de Hasbrouck (1991), suponen que la información transmitida por la transacción es incorporada inmediatamente en la siguiente transacción. Los datos de los que se dispone en nuestro mercado respecto a las órdenes límite recogen todos los cambios que se producen en las mismos sin distinguir si el efecto ha sido causado por la ejecución de una orden de mercado o por la introducción o cancelación de órdenes límite. A nuestro entender, un cambio en las cotizaciones que recoja información transmitida en la transacción anterior ha de suponer introducción, o incluso cancelación, de órdenes límite, es decir, reacción por parte del mercado. En este sentido o los datos son depurados correctamente, o los modelos que basan su estimación en el cambio del punto medio de la horquilla serían poco precisos en nuestro mercado.

En cualquier caso, nos parece que el modelo de Hasbrouck (1991) presenta una virtud respecto al resto ya que con su modelización permite la incorporación paulatina de la in-

<sup>16</sup> Esto sucede siempre que la orden de mercado transaccionada barra más de un nivel del libro de órdenes

formación a las cotizaciones y por tanto a los precios. En el resto de modelos el problema señalado se subsanaría en parte si se trabajase con distintos retardos a la hora de medir los cambios.

Como alternativa a las estimaciones del componente de selección adversa de los modelos descritos para medir la información asimétrica, surgen, además de las medidas tradicionales ya señaladas por las finanzas corporativas, modelos estadísticos que tratan de estimar directamente la probabilidad de que el creador de mercado se enfrente a un agente con información privada. En esta línea encontramos los modelos sugeridos por Easley *et al* (1996) y Nyholm (2002, 2003). El primero de ellos se basa en el desequilibrio entre compras y ventas en un determinado momento, mientras que el segundo, y de forma similar a éstos que descomponen la horquilla, se basa en el procedimiento de caracterización de la orden negociada. Estos modelos, concretamente el primero de ellos, es incluido en el análisis comparativo de diferentes metodologías efectuado por Clarke y Shastri (2001), mostrando peores resultados al considerar la relación de sus estimaciones con las variables *proxys* de la información privada en finanzas corporativas, y una correlación negativa con las estimaciones de los modelos descritos en este trabajo.

Por último, debemos señalar que, a pesar de la dificultad que entraña trabajar con las bases de datos intradiarias, consideramos de vital importancia dentro de los estudios referidos a finanzas corporativas el uso y la introducción de estos modelos a la hora de estudiar el fenómeno de la información privada. Lo que si sería conveniente a nuestro entender es confirmar la robustez de los resultados utilizando metodologías alternativas. Por otra parte, y a diferencia de los trabajos señalados en el párrafo anterior, sí que creemos que la descomposición de la horquilla en sus distintos componentes puede aportar información relevante respecto al uso de la horquilla como tal a la hora de estudiar la negociación informada.

#### 4. CONSIDERACIONES FINALES

En este trabajo hemos repasado la construcción, métodos de estimación y resultados de un conjunto de modelos cuyo objetivo común es la estimación del componente de selección adversa junto con los otros componentes de la horquilla de precios cotizada a partir de la información inferida de las cotizaciones y transacciones realizadas en las distintas jornadas de negociación en los mercados.<sup>17</sup>

Frente al conjunto de modelos que con el mismo objetivo basan sus estimaciones en las propiedades estadísticas de las series de precios, más concretamente en las covarianzas de las series de rendimientos,<sup>18</sup> el conjunto de los modelos aquí revisados se construye a partir del proceso de formación de precios atendiendo al proceso por el cual la información se incorpora a los precios de los activos.

Aunque la disponibilidad y complejidad de los datos que requieren este conjunto de modelos para realizar las diferentes estimaciones frente a aquellos otros basados en las covarianzas es mayor, la caracterización de los modelos aquí recogidos los hace especialmente atractivos para el estudio de la dinámica de formación de precios en los mercados

<sup>17</sup> Véase Clarke y Shastri (2001)

<sup>18</sup> Dentro de este grupo de trabajos podemos señalar el de Roll (1984), Stoll (1989), entre otros.

y del contenido informativo de las transacciones. Por otra parte, y como muestra la evidencia empírica reciente, estos modelos frente a los planteados en función de las covarianzas, son más precisos a la hora de diferenciar el efecto que características de la empresa como son el tamaño o el volumen contratado, tienen en la importancia del componente de selección adversa en el total de la horquilla.

Los resultados de los trabajos revisados muestran la existencia de un componente de selección adversa positivo y significativo para todos los mercados analizados y para las diferentes muestras analizadas en todos ellos. Esto significa que, en contra de los supuestos de la eficiencia del mercado, parece que las distintas transacciones tienen contenido informativo y provocan ajustes en los precios cotizados.

En lo que se refiere a la aplicación de estos modelos en otros trabajos [Ranaldo (2002) y Menyah y Paudyal (2000)], la evidencia muestra, además de resultados dispares, resultados que no están de acuerdo con las previsiones y que dependen de las hipótesis que subyacen al modelo. Por ello, la investigación en este campo, no solo la encaminada al uso de los conceptos que se desprenden de los modelos analizados sino también a analizar la validez de las medidas propuestas y la comparación entre las mismas, es creciente. Además, la mayoría de las hipótesis bajo las cuales se desarrollan los modelos están basadas en el funcionamiento de los mercados dirigidos por precios, por lo que sería interesante analizar la validez de los mismos en mercados dirigidos por órdenes como el español y/o desarrollar modelos específicos para este tipo de mercados.

En cualquier caso, el uso de estos modelos para medir la información asimétrica es de gran utilidad en el estudio del proceso de formación de precios en general y en el campo de las finanzas corporativas en particular siempre y cuando se tenga en cuenta las limitaciones señaladas a la hora de interpretar los resultados alcanzados e incluso se compruebe la robustez de los mismos con el uso de más de una de las metodologías planteadas.

## **BIBLIOGRAFÍA**

- AMIHUD, Y. y H. MENDELSON (1980): «Dealership market: market making with inventory», *Journal of Financial Economics* 8, 31-53.
- BAGEHOT, W. (1971): «The only game in town», *Financial Analyst Journal* 10-2
- CLARKE, J y K SHAOSTRI (2001): «On information asymmetry metrics», *Workig Paper SSRN*.
- COPELAND, T. y H. STOLL (1990): *Trading Markets* en Handbook of Modern Finance, 2.<sup>a</sup> ed., Warren, Gorham & Lamont. Boston.
- COPELAND, T.E. y D. GALAI (1983): «Information effects on the bid/ask spread», *Journal of Finance* 38, 1257-1269.
- DEMSETZ, H. (1968): «The cost of transacting», *Quarterly Journal of Economics* 82, 33-53.
- EASLEY, D. y M. O'HARA (1987): «Price, trade size, and information in securities markets», *Journal of Finance* 51, 1205-1236.
- EASLEY, D; N. KIEFER; M. O'HARA y J. PAPERMAN (1996): «Liquidity, information, and infrequently traded stocks», *Journal of Financial Economics* 19, 69-90.
- GLOSTEN, R y L. HARRIS (1988): «Estimating the components of the bid-ask spread» *Journal of Financial Economics* 21, 103-122.
- GLOSTEN, L. y P. MILGRON (1985): «Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders», *Journal of Financial Economics* 12, 71-100.
- HASBROUCK, J (1991): «Measuring the information content of stock trades», *Journal of Finance* 46, 179-207.

- HASBROUCK, J (1991): «The summary informativeness of stock traders: an econometric analysis», *Review of Financial Studies* 4, 571-595.
- HO, T. y H. STOLL (1981): «Optimal dealer pricing under transactions and return uncertainty», *Journal of Financial Economics* 9, 47-63.
- HUANG, R. y H. STOLL (1997): «The components of the bid-ask spread: a general approach», *Review of Financial Studies* 10, 995-1034.
- IBÁÑEZ, A.M. (2001): «Incorporación de la información en el Mercado Bursátil Español: Impacto del anuncio de una OPA». Tesis Doctoral. Universidad de Valencia.
- KIM, S.H. y J.P. OGDEN (1996): «Determinants of the components of bid-ask spread on stocks», *European Financial Management* 1, 107-125.
- KYLE, A.P. (1985): «Continuous auctions and insider trading», *Econometrica* 53, 1113-1135.
- KYLE, A.P. (1989): «Informed speculation with imperfect competition», *Review of Economics Studies* 56, 317-356.
- LIN, J; G. SANGER y G. BOOTH (1995): «Trade size and components of the bid-ask spread», *Review of Financial Studies* 8, 1133-1203.
- MADHAVAN, A; M. RICHARDSON y M. ROOMANS (1997): «Why do security prices change? A transaction-level analysis of NYSE stocks», *The Review of Financial Studies* 10, 1035-1064.
- MENYAH, K. y K. PAUDYAL (2000): «The components of bid-ask spreads on the London Stock Exchange», *Journal of banking and Finance* 24, 1767-1785.
- NYHOLM, K. (2002): «Estimating the probability of informed trading», *Journal of Financial Research* 35, 485-505.
- RANALDO, A (2002): «Market dynamics around public information arrivals» *FAME*, research paper 45.
- ROLL, R. (1984): «A simple implicit measure of the effective bid-ask spread in an efficient market», *Journal of Finance* 39, 1107-1119.
- STOLL, H.R. (1989): «Inferring the components of the bid-ask spread: Theory and empirical tests», *Journal of Finance* 44, 113-112.
- VAN NESS, B; R. VAN NESS y R. WARR (2001): «How well do adverse selection components measure adverse selection?», *Financial Management*, Otoño, 77-98.